

LOS CAMINOS DE LA EVALUACIÓN DE POLÍTICAS PÚBLICAS: UNA REVISIÓN DEL ENFOQUE

Ramón Bouzas Lorenzo

*Universidade de Santiago de Compostela
Departamento de Ciencia Política e da Administración*

Este trabajo pretende trazar un mapa contributivo de las aportaciones teóricas al campo de la evaluación de políticas públicas e, igualmente, resaltar sus debates más destacados. A lo largo de estas páginas se explora el sentido del desarrollo del campo de la evaluación de políticas, el debate sobre su función y utilidad, así como las diferentes perspectivas que han correlacionado la elección de determinadas estrategias metodológicas y la finalidad de la acción evaluativa.

***Palabras clave:** Políticas públicas, evaluación, Administraciones públicas, sector público, gestión pública.*

This paper aims to map the theoretical contributions to the field of public policy evaluation. Furthermore, it tries to highlight its outstanding debates. Both the meaning of the development of the field and the discussion on its function and utility are explored. Finally, the different approaches correlating the choice of certain methodological strategies and the target of evaluation are dealt with.

***Keywords:** Public policy, evaluation, public Administration, public sector, public management.*

INTRODUCCIÓN

La complejidad del estudio de las políticas públicas está ilustrada por la diversidad de modelos elaborados con el fin de aclarar el enorme caudal de mecanismos y procedimientos que contienen los procesos de decisión públicos. Entre las interpretaciones realizadas, nacidas de la necesidad de fijar una suerte de orden conceptual que permitiese abarcar el estudio de las acciones que intervienen en la gestación y puesta en práctica de las políticas y de evitar su aparente oscuridad y complicación, las concepciones cíclicas del *policy making* han pretendido simplificar y racio-

nalizar los procesos de toma de decisiones asumiendo la articulación de las políticas a través de distintas etapas¹, que culminarían en la puesta en práctica de diversas técnicas de evaluación, que precederían, a su vez, a una terminación o redefinición de la acción promovida: a partir de la deliberación sobre la oportunidad de resolver un determinado problema, la elaboración de una política afrontaría diferentes planos decisionales que abarcarían desde la determinación del diseño mismo de la intervención hasta el cuestionamiento del proceso completo de ejecución, la verificación del ajuste de los resultados a las in-

tenciones iniciales y la consideración de la persistencia futura de la política.

A pesar de que repetidamente se han puesto en evidencia las debilidades de un esfuerzo teórico descriptivo que en más ocasiones de las deseadas ha dejado patente su desajuste con la realidad², las visiones cíclicas del *policy-making*, integradas en el paradigma del análisis racional³, además de haber ofrecido una guía valiosa para mitigar la amenaza de dispersión y error interpretativo a los que el análisis puede conducir, han contribuido a resaltar una concepción transversal de la evaluación en el contexto del sistema político-administrativo⁴, que con el paso del tiempo, por su difusión, aceptación y utilidad, proporcionaría una notable popularidad a las prácticas *ex-ante*; a la monitorización o evaluación *in-itinere*; y a las técnicas *ex-post*.

Aunque la evaluación de políticas ha resultado ser una parcela de investigación muy sensible a los cambios de paradigma que se han sucedido en el campo general de las políticas públicas, si se considera el peso que tuvo el análisis racional durante la segunda mitad del siglo pasado y el papel que este paradigma reservó a la evaluación en el marco del examen del *policy making*, se podría llegar a afirmar, sin temor a exagerar, que el efecto que surten algunos de los éxitos y fracasos que jalonan la evolución de la teoría y práctica de la evaluación habría contribuido a suscitar, en última instancia, algunos de los altibajos que en términos de interés, popularidad y productividad científica sufriría el estudio de las políticas públicas entre la más amplia comunidad de politólogos durante los últimos treinta años. Si todavía profundizamos más en la trayectoria histórica, podemos llegar a constatar que estamos ante el primer campo que se aleja de la disciplina del análisis de políticas públicas al someter a discusión algunos temas fundamentales: la idea de un juicio técnico tanto más competente cuanto más externo resulta respecto a las dinámicas entre

los diferentes intereses y las visiones divergentes de los problemas; la representación del proceso de políticas públicas como un desarrollo lineal que discurre desde los problemas a las soluciones; la separación clara entre la identificación de los fines y la selección de los medios; y, entre otros aspectos controvertidos, la remisión reiterada al empleo de técnicas cuantitativas.

Con el ánimo de complementar algunos análisis precedentes brillantemente resueltos por otros colegas (BUSTELO, 2003; BALLART, 1998), este trabajo intenta trazar un mapa contributivo de las aportaciones teóricas al campo de la evaluación de políticas públicas e, igualmente, resaltar los debates más destacados que han tenido como telón de fondo una de las áreas de la ciencia política y de la administración que durante las últimas décadas ha mostrado mayor productividad e impacto en la práctica profesional.

DESARROLLO DEL CAMPO

La abundante literatura sobre políticas públicas y, en concreto, la que se preocupa por sus vías de análisis, nos recuerda sistemáticamente la categoría que cobra el subcampo de la evaluación a medida que crece la importancia de la intervención estatal en los países industrializados a partir de finales de los años cincuenta y principios de los sesenta, y posteriormente, durante la década de los noventa, en un evidente crecimiento paralelo a la difusión de las ideas del nuevo gerencialismo público de las que en la actualidad se siguen nutriendo académicos y profesionales.

Frente a la dilatada historia que envuelve las prácticas evaluativas que podríamos calificar de seudocientíficas, de recalitrante vigencia en sus aplicaciones al ámbito público y al privado y carentes en la misma proporción tanto de refinamiento técnico y estricta vigilancia metodológica como de sustrato ético, la evaluación que llega a merecer un tratamiento exclusivo

en el mundo académico y que aquí nos ocupa, menos dada a la espontaneidad, y más activa, consistente, sistemática y completa en lo que atañe a su aplicabilidad al proceso completo de políticas públicas, encuentra sus raíces más profundas en un segmento temporal que no supera los cuarenta años⁵.

Deteniéndonos brevemente en la evolución del campo, el estudio de la literatura estadounidense sobre la evaluación de programas y políticas de las últimas cuatro décadas revela una trayectoria retratable en tres etapas, de impulso, declive y resurgimiento⁶.

El periodo de impulso de la evaluación abarca la década de los sesenta y se prolonga hasta bien entrados los setenta. Se trata de un momento en el que las fuertes inversiones en materia de evaluación, amparadas por una favorable iniciativa legislativa, presupuestaria e institucional⁷ y por una desbordante euforia entre los académicos sobre la utilidad práctica de las ciencias sociales⁸ y su facultad para ofrecer respuestas acerca del rendimiento de los programas y sobre la medida en que éstos contribuían al éxito de las políticas, favorecen la formación y consolidación de amplias comunidades de investigadores, la definición de los límites científicos del área, y el fomento, interno o externo, de unidades y organismos responsabilizados de esta materia⁹.

Factores de índole económica, combinados con un importante cambio en la ideología de la gestión pública, hicieron que desde mediados de los setenta hasta el inicio de los noventa el interés por la evaluación decreciese. Las unidades responsabilizadas de la evaluación, consideradas en algunos casos, por sus funciones de control, una amenaza –o, cuando menos, una fuente de malas noticias– para los clientes internos, no llegaron a gozar en muchas agencias gubernamentales de la popularidad suficiente para garantizarse su supervivencia en periodos de intensos

recortes de los recursos disponibles. Este fenómeno, por sí sólo poderoso como explicación del declive del cultivo de la evaluación, unido a una defensa extrema de la reducción del sector público, inspirada en las ideas de la nueva gestión pública, pudo haber contribuido decisivamente a acelerar un cambio en las preferencias profesionales de los académicos que haría disminuir el interés por esta disciplina y obligaría a los autores más firmemente convencidos de las bondades de la evaluación a replegarse en el debate teórico en detrimento de la investigación más empírica y el desarrollo y perfeccionamiento tecnológico.

Desde mediados de los noventa, un nuevo impulso legislativo¹⁰, una renovada confianza en las capacidades operativas de la Administración y la creencia –cada vez más compartida por las elites¹¹– de que, con independencia de los recursos económicos, a través de una adecuada formación (cultural-profesional) de los empleados y una concepción más flexible y esmerada de los procesos administrativos y del diseño estructural¹² se podría obtener un desempeño eficiente, un elevado rendimiento y una gestión de calidad acorde con las expectativas de la ciudadanía, logran situar nuevamente en un primer plano el interés por la evaluación¹³. Las reformas realizadas a finales de los noventa en los Estados Unidos en diversas políticas sociales y, en particular, en educación, serían testigos de la introducción de un conjunto amplio de técnicas en las que, frente a las modalidades verticales-descendentes vigentes durante los sesenta, las evaluaciones de fin de siglo, con un tono más pluralista, introducen mayores mecanismos de participación y adquieren una mayor orientación al usuario-destinatario.

Como antes se ha señalado, estos tres periodos, además de segmentar el cultivo de la evaluación e iluminar los factores de índole contextual que han dejado mayor huella en la ganancia o pérdida de popularidad del campo, nos ayudan, en un plano

superior, a tomar el pulso a la investigación más extensa sobre políticas públicas. Igualmente, en lo que respecta a la evolución de este terreno en otros entornos, con matices de importancia variable según el área territorial a considerar, la pauta descrita puede servir para esbozar la trayectoria de los estudios sobre evaluación en el ámbito europeo (TOULEMONDE, LEEUW y BROUWERS, 1999; FONTAINE y MONNIER, 1998; TOULEMONDE, 1995): así, aunque en general el campo comienza a desarrollarse a partir de los setenta, mientras en el norte del continente y en el área de influencia anglosajona la apreciación del valor que las técnicas de evaluación pueden proporcionar a las organizaciones públicas tiene reflejo inmediato en la práctica¹⁴, en el resto del continente el mayor peso recaería en una producción teórica de desigual profundidad –en ocasiones, no demasiado alejada de lo que podría considerarse literatura introductoria, con propósito meramente divulgativo-, carente de la proyección práctica sólida y rigurosa¹⁵ que hubiese sido necesaria para dotar de credibilidad desde un principio al estudio de la evaluación de políticas entre académicos y gestores y evitar que factores como la reducción del déficit público o la lucha contra la corrupción se convirtiesen, casi en exclusiva, en lo que parecían ser las utilidades y argumentos determinantes para animar a muchos politólogos a reorientar sus intereses hacia esta parcela de las políticas públicas.

PERSPECTIVAS SOBRE LA FUNCIÓN Y UTILIDAD DE LA EVALUACIÓN

Dejando a un lado la crónica sobre la consolidación del campo o, en el caso español, la del modesto recorrido –a pesar del esfuerzo realizado en la formación de grupos de investigación- en pos de la conquista de la difícil misión de crear una tradición científica, aspectos que ya han suscitado la inclusión de innumerables comentarios en las principales obras de la

disciplina, podemos ahora detener la mirada en otro punto debatido prolongada y abundantemente por los principales teóricos del área: la consideración de las funciones y de la utilidad de la evaluación.

En torno a las funciones y a la utilidad, volviendo al marco del proceso de elaboración de políticas públicas, la relación directa de la evaluación con la toma de decisiones, fuera de toda duda, ha sido considerada desde dos planos: en primer lugar, la visión más tradicional ha resaltado la evaluación como un proceso cardinal que ofrece soporte al proceso de toma de decisiones al proporcionar valoraciones del grado en que los programas y políticas diseñados han logrado ajustarse a las metas propuestas inicialmente¹⁶; y, en segundo lugar, partiendo de una comprensión más rica, tanto teórica como prácticamente (potencialidad de la variedad técnica), la evaluación es considerada una función de gestión con capacidad para aportar un valor añadido a la actividad administrativa al ofrecer soluciones tecnológicas para determinar con fines formativos o sumativos no sólo el nivel de logro y percepción de los resultados finales en relación con los objetivos previstos inicialmente, sino también la generación de objetivos imprevistos; el examen de los efectos sobre la población objeto y el entorno; la verificación de la consecución de los objetivos en relación a la asignación óptima de recursos; o la evaluación monetaria del objeto (coste-beneficio), de sus costes y efectividad (coste-efectividad), así como de sus resultados en relación a unas categorías de utilidad (coste-utilidad)¹⁷ (*vid.* gráfico 1).

En cierto modo, en términos más familiares, las dos ideas destacadas guardarían relación con la distinción de las funciones elementales que STUFFLEBEAM y SHINKFIELD (1995) atribuyen a la evaluación: la de perfeccionamiento (*improvement*), desde la que se concibe la evaluación como un sistema con capacidad para dar respuesta a las necesidades gerenciales, garantizar unas prestaciones óptimas

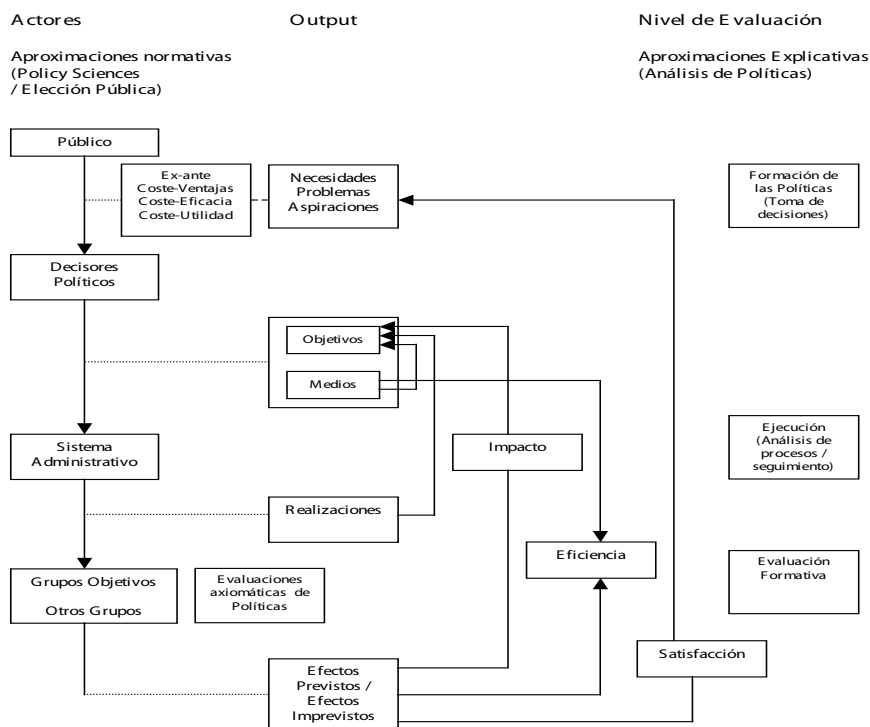


Gráfico 1: *El proceso de políticas públicas y las modalidades de evaluación*. Fuente: Adaptación de NIOCHE (1982).

y proporcionar, simultáneamente, recursos para la mejora continua de la propia acción evaluativa; la rendición de cuentas (*accountability*), con la que se pretende responder al compromiso básico de satisfacción de la necesidad de clarificar y hacer perceptible y comprensible a la ciudadanía la acción de gobierno y el rendimiento de la Administración en los sistemas democráticos¹⁸; y la guía de acciones futuras (*enlightenment*), desde la que se ofrece un soporte para el aprendizaje organizativo, al facilitar unos medios y una tecnología ya contrastados ante el posible diseño de nuevas acciones¹⁹.

Con el fin de proporcionar soluciones ante la extensa gama de intervenciones evaluativas que desde la perspectiva de

la *accountability* deben poner en marcha los diferentes actores institucionales (partidos; parlamentos; tribunales de cuentas, etc.), la literatura se ha caracterizado por su remisión a la metodología de la investigación en ciencias sociales²⁰ para buscar técnicas y modalidades que permiten ofrecer nuevas formas de calibrar el rendimiento de las políticas y de los programas específicos de actuación pública. En este sentido, desde la limitada concepción de la evaluación como ejercicio elemental de verificación de resultados y contraste con los objetivos iniciales hasta la percepción de su potencial riqueza para perfeccionar y dar sentido –desde un punto de vista finalista– a la acción política, ha crecido toda una amplia gama de campos de aplicación

que comprenden la práctica extensión del dominio del análisis de las políticas públicas. Ha sido esta exploración de las estrategias de análisis y de la diversidad de los terrenos que la actividad evaluadora ha abarcado lo que ha convertido esta área en una de las más agitadas ágoras pluridisciplinarias de la ciencia política.

A partir de la paráfrasis de la definición que ofrece M. Scriven sobre la evaluación²¹, E. Vedung, en uno de los manuales de más reciente éxito en este campo, recapitula las prácticas de evaluación y los contextos en los que se llevan a cabo, reclamando atención sobre diversas observaciones metodológicas. Elabora, en el fondo, lo que podría ser interpretado como un elenco de los temas que más debate han generado al teorizar sobre la práctica evaluativa:

Intervenciones gubernamentales: aunque como proceso analítico la evaluación puede ser empleada en cualquier terreno de las ciencias sociales, es indiscutible su carácter político²² y su identificación con la actividad gubernamental, donde es aplicable tanto al área de las grandes líneas de intervención pública (políticas públicas) como a los procesos que afectan transversalmente a la operatividad de la Administración (diseño estructural; diseño de procesos; planificación; control; presupuestación, etc.)²³.

Mecanismos de procesamiento, productos e impactos: en el seno de la actividad de cualquier organismo los procedimientos de evaluación no sólo atienden a los métodos de procesamiento de insumos, sino también a los productos que la transformación genera y a los impactos que a corto, medio y largo plazo suscita la actividad.

Valoración de actividades en curso y finalizadas: la evaluación orientada a las actividades en curso (ajuste a criterios de eficiencia)²⁴ no excluye, desde la perspectiva de la creación de vías de aprendizaje y perfeccionamiento, el examen *ex-post* de las intervenciones conclusas.

Retrospección y anticipación: la evaluación implica una valoración retrospectiva de las intervenciones públicas e, igualmente, comporta mucho más que la mera apreciación de lo activo; facilita la puesta en marcha de mecanismos susceptibles de fundamentar los procesos de debate y la toma de decisiones²⁵ relativos a programas y políticas venideros²⁶.

Examen de la eficiencia, administración y supervisión de producción: la gama de modalidades de evaluación se extiende más allá de la clásica evaluación de impacto; además de comprender el estudio de la eficiencia, la administración y el seguimiento²⁷, atañe a la calidad de las prestaciones.

Diversidad de criterios de valoración: la evaluación no sólo se limita a la verificación del cumplimiento de objetivos, sino que abarca la apreciación de muchos otros elementos de tangibilidad variable²⁸ (méritos, valor monetario, valor moral) que ofrecen interés para quienes directa o indirectamente –diseñadores de la intervención, promotores, mediadores, agentes críticos (*stakeholders*), receptores– están involucrados en las acciones emprendidas.

Empleo de métodos formales y rigurosos: la evaluación contemporánea, en especial, la llevada a cabo en el ámbito público, trata de ajustarse metodológicamente a unos mínimos criterios de calidad y rigor y abandonar prácticas pseudocientíficas, escasamente solventes, nacidas del oportunismo, la improvisación o la intuición.

LITERATURA Y DEBATE METODOLÓGICO

El lector que profundiza en la literatura sobre la evaluación de políticas pronto advierte que esta parcela científica ha ido creciendo, aumentando su diversidad temática, variedad metodológica y nómina de contribuyentes, al tiempo que se han intensificado los debates en torno a los fines, la metodología y las prácticas que

abarca. Hasta cierto punto, un balance riguroso de la teoría sobre la evaluación, además de difuminar las diferencias detectables entre las diversas perspectivas, impide, en sentido estricto, distinguir entre aproximaciones fundacionales y actuales: lo que se podría considerar inaugural, lejos de caer en desuso y alejarse de la práctica contemporánea, está muy presente todavía en los diferentes sectores en los que habitualmente se aplican las técnicas de evaluación²⁹.

El debate actual entre los teóricos de la evaluación parece volcarse más en el cuestionamiento de los paradigmas epistemológicos³⁰ que en la consideración de la función de la evaluación en el marco de las teorías sobre la elaboración de políticas públicas. En este sentido, las aportaciones siguen divididas entre aquellas que muestran su adhesión al paradigma experimental, caracterizado por el alineamiento con el positivismo lógico, la metodología cuantitativa³¹ y el empleo de cadenas deductivas y de generalización; y las que se identifican más con el constructivismo social y, en consecuencia, optan por una epistemología postpositivista, adoptada tanto desde un enfoque hermenéutico postmoderno –centrado en la elaboración de análisis holísticos y la exaltación de los métodos cualitativos³²– como desde un enfoque tecnológico o pragmático –donde se podrían situar las aportaciones de D.B.Fishman.

Con afán recapitulatorio del debate sobre la metodología y la función de la evaluación, distanciándonos de la literatura sobre la evaluación nacida de las teorías del análisis de políticas públicas, y prescindiendo del empleo de un criterio cronológico en el intento de desentrañar las vías abiertas en este campo, podríamos empezar por “rescatar” de las obras de D.T.Campbell y M.Scriven una primera aproximación al estudio de la evaluación de políticas.

Ambos autores son exponentes de un enfoque “puro” en el que, desterrada

toda influencia de los procesos decisionales, la evaluación es estudiada como análisis de los resultados de las políticas: los evaluadores consideran que para diagnosticar los resultados es necesario formular una hipótesis de impacto de la variable independiente (la política)³³ en la variable dependiente (los cambios producidos en aquella parte de la sociedad sobre la que ha actuado la política pública). Con el fin de profundizar en esa cuestión, en una de las monografías más destacadas de la disciplina, CAMPBELL y STANLEY (1963) teorizan sobre diseños de investigación experimentales y cuasi-experimentales, y disertan sobre los modelos causales de impacto como solución para un problema al que desde muy temprano se enfrentaron los evaluadores³⁴: la validez de las relaciones de impacto. Este se convertiría en un tema de debate crucial hasta el punto de que las diversas nociones de validez³⁵, especialmente la interna, inspirarían algunos instrumentos de evaluación célebres nacidos del refinamiento econométrico. Se trataría, en suma, de considerar la potencialidad del control experimental y cuasi-experimental para el análisis de las políticas públicas³⁶.

Alineada con esta primera perspectiva, en la obra de M.Scriven y, concretamente, en su consideración de la lógica de la evaluación³⁷, se podría hallar un buen testimonio de la profunda creencia en la potencialidad y riqueza de la práctica de la evaluación hasta el punto de minusvalorar el análisis de los procesos de toma de decisiones y soslayar el debate sobre la consideración de la función de la evaluación en el marco más amplio del estudio de las políticas públicas. Más allá de sus aportaciones sobre la evaluación formativa y sumativa, por las que es ampliamente reconocido, Scriven defendería la edificación de una ciencia de la evaluación justificada –con beneficios superiores a los costes que comporta³⁸–, orientada al interés general³⁹ y verdaderamente comprometida con la necesidad que la ciudadanía tiene de co-

nocer las características de la actuación de quienes se responsabilizan de su administración. Para ello, llegaría a proponer un modelo evaluativo ambicioso, en el que además de superar las constricciones de las concepciones más racionales, basadas en la importancia de la consecución de los objetivos –asumiendo así las tesis de los defensores de la controvertida, aunque revolucionaria, concepción de la evaluación libre de objetivos (*goal-free evaluation*) - se pondrían en tela de juicio los procedimientos mismos de evaluación (metaevaluación).

Las aportaciones de esta primera corriente, aunque muy cuestionadas durante los últimos veinte años (ROSSI y FREEMAN, 1989), se mantienen muy vivas: en términos de debate, a pesar de las críticas recibidas –casi todas basadas en aspectos epistemológicos-, han contribuido a cuestionar el papel de la evaluación tanto desde la perspectiva del proceso de gobierno de las políticas públicas como desde el plano del uso del conocimiento; y, en términos metodológicos, aun considerando su limitación y lentitud⁴⁰, se sigue reconociendo la idea de que el control causal –experimental o cuasiexperimental- puede ser utilizado con éxito en variados sectores de las políticas públicas.

La reacción a esta primera línea podríamos encontrarla en varios enfoques diferenciados.

En primer lugar, frente a la visión ingenuamente aséptica del papel desempeñado por el evaluador ante los procesos de toma de decisiones, propia del enfoque adoptado por Scriven y Campbell, hay autores que, como WHOLEY (1983), contraponen la figura de un evaluador que opera en estrecho contacto con los responsables-gestores de los programas y proponen una concepción en la que la evaluación se transforma en un instrumento que permite el logro de una gestión más eficiente, un mayor rendimiento y mejores resultados⁴¹: la evaluación, en tanto que herra-

mienta de gestión, flexible, adaptable a todo tipo de situaciones, se amolda a las características de los programas y políticas emprendidos.

Desde esta línea, la comunidad académica y profesional dedicada a la evaluación debe adoptar una actitud mucho más comprometida con la misión que persiguen los directivos públicos y ser consciente de la disponibilidad de una amplia gama de técnicas que pueden permitir eliminar o atenuar la influencia de algunos factores que pueden malograr la evaluación: coste de recogida de datos; variabilidad de las actividades a examinar; nivel administrativo y ámbito territorial en el que opera; actitud y motivación de los actores implicados; o tipo de objetivos propuestos y claridad de las hipótesis formuladas. Con el fin de enriquecer la función del evaluador y proyectar sobre el terreno práctico la necesidad de refinamiento técnico, la literatura ofrecería pautas de actuación encaminadas a optimizar el empleo de los instrumentos de evaluación⁴² y prescripciones sobre la manera de gestionar el diseño de las intervenciones, con especial atención a las interacciones a las que dan lugar en el ámbito político los programas y las políticas públicas⁴³.

En una posición reivindicativa de la defensa de la evaluación y su papel social muy semejante a la de Wholey, pero, si cabe, con una pretensión menos elitista, menos orientada al cumplimiento con el directivo público y más comprometida con la satisfacción de los destinatarios reales de las intervenciones, se situaría la contribución de R.E.Stake, quien reubicaría nuevamente la función del evaluador –centrada en la prestación de servicios- y reinterpretaría de una manera más pragmática el empleo de las técnicas –de marcado carácter cualitativo-, subordinando el rigor a la utilidad⁴⁴. Desde esta óptica, la evaluación pierde parte de su tono más exclusivamente tecnocrático, y adquiere un carácter dialogante y abierto, rasgos que son contemplados como condición de

una contribución a una gestión más eficaz y más democrática de los servicios públicos (NIOCHE, 1982: 56).

Aunque respecto a la visión de Scriven y Campbell, tanto desde la perspectiva de Wholey como la de Stake, persiste el peso de una teoría racional-sinóptica del proceso de elaboración de políticas públicas, las concepciones del saber y de la función del evaluador, así como los diseños de investigación –a excepción de la evaluación intensiva, en la que se evidencian los controles clásicos– se modifican parcialmente. Desde este punto de vista, los procesos de toma de decisiones, la función del evaluador y el aparato teórico-metodológico pasan a formar parte de un todo. Se trata, en otras palabras, de una visión de la evaluación fundamentada en la primacía de la asunción social de las ideas por encima de su certeza⁴⁵.

Hay, no obstante, otras vías para reposicionar la evaluación de políticas: en la misma insatisfacción ante el modo de concebir la toma de decisiones y la función del saber científico en los procesos políticos se encuentra el elemento que explica el alejamiento gradual de muchos teóricos de la evaluación respecto a los planteamientos originales. En este sentido, C.Weiss nos recuerda que, aunque a menudo resulte inadvertido por los autores menos sensibilizados con los procesos decisionales, muchos programas y políticas con buen rendimiento son una consecuencia del éxito de decisiones políticas precedentes, y que, de hecho, la confrontación política en torno a los programas no concluye tras la aprobación inicial de la intervención⁴⁶. En este orden de cosas, considerando la naturaleza del proceso que conduce a la fijación de los objetivos, el carácter estratégico que adopta la conducta de los diferentes actores involucrados e, incluso, el papel que desempeña el científico social en las políticas públicas, es imposible dejar de concebir la evaluación de políticas como una actividad intrínsecamente política⁴⁷. Así, en la propia consideración de la existencia

de una política de la evaluación radicaría la razón ser de diferentes manifestaciones del uso de la información: el empleo manipulativo de la investigación social en función de intereses específicos de grupos de presión; la producción de información por motivos de control de la organización o del ambiente; o la sobresaturación de información respecto los recursos necesarios para la toma de decisiones⁴⁸.

Con el fin de superar las limitaciones que se derivan de la intervención de factores de índole política en los modelos simplistas de resolución de problemas y frente al modelo racional-sinóptico, L.J. Cronbach presentaría el modelo de la *policy-shaping community*, en el que sostendría que la evaluación no se conforma respecto a normas o modelos abstractos, sino que se define a partir de las características de la comunidad de políticas que la auspicia⁴⁹. Desde esta óptica, Cronbach pondría en tela de juicio las evaluaciones que a partir del empleo de técnicas *ex-ante* aspiran a decir la última palabra sobre una política –su peligro radicaría en su tendencia a inducir a sus responsables a prestar mayor atención a los impactos, a menudo imprevisibles, y a descuidar la vigilancia de los *inputs*, más fácilmente controlables– y, en términos metodológicos, privilegiaría la amplitud de las cuestiones tratadas frente a la precisión de las relaciones causales de impacto, siendo más importante para Cronbach reconstruir el problema que la política abarca y redireccionar el diseño de las actuaciones futuras que medir las relaciones causales a partir de las experiencias precedentes⁵⁰.

En el fondo de su reflexión metodológica y teleológica, Cronbach debate nuevamente la función de la evaluación de políticas, que, según sostiene, no debería preocuparse demasiado ni de la resolución de problemas sociales ni de la utilización de un instrumental científico, sino de procurar la mejora del proceso de producción de políticas⁵¹. Y, en esa línea, aunque con un tono teórico sensiblemente diferente⁵²,

Weiss considera que la evaluación debería promover una infiltración de nuevas ideas, diferentes cuadros conceptuales y definiciones alternativas de los problemas públicos en los diseñadores de políticas (*enlightment*): en concreto, su concepción de la “ilustración” o “iluminación” se fundamentaría en la idea de que, aun en un contexto condicionado por los intereses de los diferentes actores, la ideología y la disponibilidad de información, la labor de investigación tiene una gran potencialidad para esclarecer las vías que permiten superar las complejidades que habitualmente acompañan los problemas a solventar en el ámbito público.

En este punto, en la actualidad, el debate teórico se sitúa en un aspecto muy polémico: por un lado, es loable el intento de “iluminación” de los *policy-makers* que desea Weiss y, por supuesto, es más que razonable pensar que la evaluación forma parte de un “proyecto” más ambicioso de políticas públicas que debe ser mejorado⁵³; por otro lado, como señalaría Lindblom⁵⁴, existen muchas otras sendas alternativas que pueden evitar que una concepción libre de objetivos y perfectamente incluida en el proceso de elaboración de políticas amenace con convertirse finalmente en un instrumento al servicio de las elites y de los máximos responsables de los programas, poniendo en duda el carácter democrático de todo el proceso. Tal vez, si la cuestión pendiente en la teoría de la evaluación es la de dilucidar el empleo del conocimiento que la aplicación de las diferentes técnicas genera, probablemente, no es suficiente con diferenciar entre un uso conceptual o instrumental o entre el corto, medio o largo plazo, sino que debería ser necesario indagar las diferencias entre un uso aceptable o inaceptable.

RECAPITULACIÓN

Al contextualizar los debates de la teoría de la evaluación en el campo más amplio de la ciencia política y de la ad-

ministración, DENTE y RADAELLI (1995: 147-185) subrayan tres aspectos: primero, que la eficacia de las evaluaciones depende decisivamente de las teorías de la elaboración de políticas en que se insertan; segundo, que la cuestión del impacto de las evaluaciones se está afrontando, principalmente, dentro de modelos de análisis que endogenizan el conocimiento, mientras parecen quedar en un segundo plano los modelos de examen del impacto basados en los marcos interpretativos que en el corto, medio y largo plazo estructuran los problemas de la acción política; y tercero, que la evaluación de las políticas evidencia la posibilidad de una ciencia política aplicable, basada en procesos de control diferentes a los derivados de los métodos comparados.

Siguiendo diferentes intereses –homogeneización metodológica; refinamiento técnico; examen de la función y utilidad; papel del evaluador⁵⁵ y base deontológica; sentido disciplinar e interacción ámbito profesional-político-, casi siempre interrelacionados, las variedades interpretativas de la evaluación nacidas desde los setenta, retroalimentadas por la práctica, han conseguido delimitar y legitimar un campo de estudios que, a pesar de crecer contrayendo una gran deuda con otras parcelas disciplinares de las ciencias sociales, ha alcanzado una madurez indiscutible y, sin duda, ha abierto exitosamente una nueva vía de debate teórico y de cultivo práctico en la ciencia política y de la administración.

Como indica CHELIMSKY (1997), el contexto político más global en el que emergen nuevos sectores demandantes de intervención (protección medioambiental; desarrollo sostenible; derechos humanos); el desarrollo normativo que impulsa la actividad de los evaluadores, la aparición de nuevos actores, temas y necesidades; la revitalización de problemas tradicionales relacionados con las funciones de rendición de cuentas y responsabilidades, desarrollo y mejora de programas, políticas e instituciones y conocimiento; y el aprendizaje de

nuevas lecciones extraídas de la continua experiencia en evaluación parecen garantizar, en suma, una fructífero porvenir a este campo.

NOTAS

1. La idea básica que fundamenta estos modelos es la de que la acción política surge de las interrelaciones y de las acciones de numerosos actores que entran y salen de escena, que desempeñan diversas funciones y que se plantean objetivos diferentes a lo largo de un período reductible a diferentes fases. Por medio de múltiples relaciones causa-efecto u orden-respuesta, la política se articularía en etapas, lo cual facilitaría la intervención de los analistas al intentar precisar *ex-ante* las probabilidades de éxito o fracaso o calcular *ex-post* el rendimiento de las acciones emprendidas. Se trataría, en suma, de lo que Sabatier y Jenkins-Smith (1993) concebirían como heurística de etapas.

2. Como señala Regonini (2001), en cierto modo, la evaluación nace de la constatación de que las premisas analíticas del análisis tradicional de políticas públicas no proporcionan elementos para afrontar la indeterminación de los datos y la arbitrariedad de los criterios contenidos en cada verificación de la divergencia existente entre lo que una política debería haber sido teóricamente y lo que ha resultado ser en la práctica.

3. Como recuerda Regonini (2001: 106-130), el análisis racional hace suya una terminología (eficiencia; *input*; *output*; preferencia, etc.) presente con diferentes significados en diversos paradigmas de las ciencias sociales. La explicación del éxito e impacto que entre los políticos interesados por el estudio de las políticas públicas tuvieron estos conceptos radica en la propia base teórica del *policy analysis*: economía (paretiana; del bienestar; institucional / organizativa); ciencias de la gestión (investigación y análisis operativo; teoría de sistemas; gestión pública); teoría de la toma de decisiones y teoría de la resolución de problemas.

4. Concepción ya subrayada por Lowi (1964).

5. Aunque en el ámbito estadounidense la evaluación recibe un fuerte impulso desde los sesenta en el contexto de las intervenciones de la *Great Society*, y su importancia queda refrendada hasta la actualidad en el contenido del corpus legislativo que conforma la denominada *Sunset Legislation*, en el ámbito español será

necesario esperar hasta la década de los ochenta para poder dar fe (en un ambiente dominado por los procesos de adaptación a la CEE y la puesta en marcha de diversas iniciativas orientadas a la mejora de la gestión como consecuencia de la ejecución de los programas de modernización que emprenden la Administración General del Estado y algunas Comunidades Autónomas), de que la evaluación de programas y políticas comienza a dejar de ser un simple, aunque efectivo, recurso retórico electoralista y se comienza a valorar como una herramienta de gestión imprescindible.

6. Sobre la evolución de la teoría y práctica de la evaluación en los Estados Unidos, puede consultarse el trabajo de Peters (1998). Igualmente, en la obra de Chen (1990) se hace un repaso muy interesante y completo del cambio de paradigmas en el campo de la evaluación en los Estados Unidos desde 1960.

7. Uno de los mayores hitos lo marcó el cambio de objetivos y de competencias de la vetusta *General Accounting Office*, originalmente centrada, como su propia denominación indica, en aspectos de control contable y posteriormente reorientada hacia el terreno del análisis y la evaluación. El efecto inmediato de este cambio fue la creación de una cultura de responsabilidad y la institucionalización de la función evaluativa.

8. Optimismo que llegaría a ahogar las dudas razonables que existían sobre la capacidad de extracción de información a partir de datos no experimentales (Campbell y Stanley, 1966).

9. Como señala Peters (1998: 24), "si las agencias querían tener éxito en el proceso presupuestario tenían que ser capaces de justificar que sus programas eran efectivos en el coste e incluso eficientes".

10. La producción legislativa a la que se alude, entre la que destacan la *Government Performance and Results Act* (1993) y el *Regulatory Review* (1996), refuerza la elaboración de políticas y la introducción de programas sujetos a la planificación estratégica y el examen pormenorizado de sus resultados —en el caso de la segunda norma citada, de forma prospectiva y con una orientación técnica enfocada sobre la detección de costes— en función de una gama limitada de indicadores cuantitativos.

11. En el contexto de la "reinversión del gobierno" promovida por la Administración Clinton.

12. La orientación hacia los procesos administrativos ha estimulado una de las áreas más cultivadas de la evaluación. Su éxito estaría justificado por la necesidad de verificar la operati-

vidad de un programa o una política a través de la revisión de su estructura normativa (normas, planes de acción); su estructura administrativa (agencias; comisiones; *task force*) y sus redes de comunicación, coordinación y cooperación intra e interorganizativa.

13. La celebración en noviembre de 1996 de la *International Evaluation Conference* (Vancouver, Canadá) marca un punto de inflexión en el proceso de conformación de la comunidad científica de la evaluación. Esta reunión representó la primera gran cita internacional para evaluadores, académicos, auditores y asociaciones profesionales (entre ellas la *European Evaluation Society*). Algunas de las ponencias elaboradas para esta reunión pueden consultarse en la obra de Chelimsky y Shadish (1997).

14. Si en el caso de los Estados Unidos las concepciones ideológicas de gestión dominantes durante la era Reagan reducen la potencialidad de la evaluación y convierten su tecnología en mero instrumento de control al servicio de la contención del gasto y reducción del sector público, devaluando enormemente su aceptación entre los practicantes y bloqueando significativamente su crecimiento, en el contexto europeo, con escasos antecedentes en esta materia, la evaluación encontraría un lugar entre las prácticas administrativas merced a los programas reformistas –considérese, por ejemplo, los impulsados durante el *thatcherismo*–, pero al alto precio de dar lugar a una interpretación paupérrima de su verdadero sentido gerencial y a ralentizar considerablemente su cultivo y la consecución de una madurez todavía muy discutible.

15. Como señala Bustelo (2003: 13) al comentar el cultivo teórico y práctica de la evaluación en España: “estamos todavía lejos de contar con una práctica normalizada y con una verdadera cultura de la evaluación en nuestros organismos públicos. Este todavía incipiente campo de la evaluación de programas y políticas públicas necesita, por lo tanto, de un mayor desarrollo académico y profesional. En muchas ocasiones, la falta de referencias ha significado la sustitución de una verdadera evaluación por otro tipo de prácticas de análisis que, creemos, es necesario distinguir. De esta manera, o bien se ha suplido esta falta de especialización con una aplicación estricta y aséptica de la metodología de investigación social al uso, realizando estudios parciales que han sido poco sensibles al contexto y han resultado poco útiles para las organizaciones y las políticas evaluadas, o bien se han manejado las situaciones [...] para hacer

estudios *ad-hoc* que fueran favorables a una determinada coyuntura política, dando origen a situaciones con poca credibilidad externa que han desprestigiado la práctica de la evaluación”.

16. Esta visión nos remitiría a la obra de Tyler (1967), de fuerte impacto inicial en el cultivo técnico de la disciplina y de extensa acogida en el terreno práctico, a pesar de la abundante crítica recibida durante las últimas décadas.

17. Siguiendo la terminología de Monnier (1990-91: 131-140), en esta segunda visión se encuadrarían todas las aproximaciones a la evaluación que han superado la primitiva aproximación por objetivos: aproximación por la puesta en práctica del programa (perspectiva operacional); aproximación por los efectos (eliminación de la influencia de los objetivos oficiales en las percepciones); aproximación por los procesos (interacciones del programa y el entorno) y aproximación por los actores sociales (pluralismo).

18. Comprendiendo en un sentido amplio la rendición de cuentas, Rossi y Freeman (1989) aportan seis interpretaciones de la *accountability*: legal (ajuste del programa o la política a la legislación); fiscal (ajuste de la gestión del programa a las regulaciones fiscales); distributiva (fiabilidad de la asignación de los productos o servicios producidos); de cobertura (grado de ajuste de los receptores finales con el público objetivo inicial); de impacto (coherencia del producto final con la previsión inicial); y de eficiencia (ajuste del impacto a los costes del programa). Por su parte, Vedung (1997: 104-113), considera necesario, siempre que sea posible, analizar la responsabilidad en las acciones que emprende la Administración desde cuatro perspectivas: la de las elites políticas; la de los gerentes de los organismos públicos; la de los ciudadanos y la de los clientes.

19. Aunque la obra de Wildavsky (1979) y los trabajos realizados por C.H.Weiss durante los setenta y ochenta son conocidos por su defensa de las posibilidades de aprendizaje y ulterior perfeccionamiento del rendimiento de las políticas, en las investigaciones de Hudson, Mayne y Thomlison (1992) y Preskill y Torres (1998) se pueden encontrar interesantes experiencias que confirman esa posición. Sobre la relación, en particular, entre el aprendizaje y el perfeccionamiento de las *policy communities*, consúltese a Subirats (1994).

20. Suchman (1967) justificaba el empleo de la metodología de las ciencias sociales para la determinación (basada en opiniones, documen-

tos, datos subjetivos u objetivos) de los resultados (deseables o indeseables, temporales o permanentes, inmediatos o a largo plazo) alcanzados por una actividad orientada a alcanzar un fin u objetivo. En esta misma línea se expresaba Chelimsky (1989: 7) al concebir la evaluación, por definición, como investigación social: "program evaluation is the application of systematic research methods to the assessment of program design, implementation, and effectiveness".

21. Según Scriven (1991: 1), "evaluation is the process of determining the merit, worth, and value of things". Por su parte, Vedung (1997: 3) se refiere a la evaluación como "careful retrospective assessment of the merit, worth and value of administration, output, and outcome of government interventions, which is intended to play a role in future, practical action situations".

22. En la línea de los comentarios realizados por E.Vedung, este aspecto sería resaltado en otros trabajos de teoría de la evaluación: Monnier (1990-91; 1995); Palumbo (1987) o Weiss (1988).

23. A causa de la inserción de las técnicas de evaluación en el dominio de estudios sobre los procesos político-administrativos, la literatura ha terminado por asumir, además de la producción tradicional centrada en cuestiones de método y función, los estudios sobre control de gestión, "que relacionan el campo de la evaluación con el análisis de procesos, calidad y reingeniería, y con los sistemas de información contable-financiera o de control de operaciones, propósitos de la gestión pública o privada y del análisis de organizaciones" (Ballart, 1998: 229).

24. La evaluación "en tiempo real" de las actividades pertenecientes a un programa o una política, considerada el núcleo o esencia de la labor del evaluador, ha dado lugar a toda una tradición en la literatura estadounidense: Wholey *et.al.* (1970); Nachmias (1979) y Rutman (1980).

25. En una de las visiones más clásicas sobre este particular, Anderson y Ball (1978) resaltarían la capacidad de la evaluación de programas para contribuir a tomar decisiones sobre el desarrollo, ejecución, continuación, expansión, modificación o finalización de un programa; obtener evidencias para apoyarlo o rechazarlo; y comprender procesos psicológicos y sociales básicos.

26. En la literatura son detectables posiciones muy diferentes en torno al alcance de la evaluación: Vedung (1997: 7) considera la evaluación exclusivamente retrospectiva y rechaza

admitir dentro de la especialidad toda forma de análisis tendente a la valoración prospectiva; para otros autores, sin embargo, los estudios de anticipación, identificados con la metodología *ex-ante* (evaluación de previsión; inventario de necesidades; fijación de objetivos) requieren el empleo de procedimientos de análisis, que, por su sentido y recursos tecnológicos pueden ser incluidos dentro del amplio conjunto de técnicas evaluativas (Rossi y Freeman, 1989; Anderson y Ball, 1978).

27. Aunque Vedung (1997) considera el seguimiento como una faceta más de la gestión en la que se pueden aplicar técnicas de evaluación, la literatura pone en evidencia una verdadera división de opiniones sobre la consideración de las acciones de seguimiento como actividades de evaluación. La diferencia radicaría en el carácter enjuiciatorio de las técnicas empleadas: en caso de estar presente, hablaríamos de evaluación, y, en caso contrario, de una actividad convencional de verificación / confirmación del cumplimiento de un programa. Sobre la utilidad y metodología del seguimiento, consúltese Banco Mundial (1996).

28. En contraste con las formulaciones iniciales de la teoría de la evaluación, prescriptoras del rechazo de toda subjetividad que pudiese interferir en la evaluación, la visión actual sitúa a los valores en el punto central del proceso de evaluación (Stufflebeam y Shinkfield, 1995; Shadish, Cook y Levinton, 1991).

29. La consideración misma del diseño de la evaluación –marco para la planificación y dirección del estudio– testimonia la actualidad de algunas aportaciones muy consolidadas. Sobre este aspecto es útil revisar las aportaciones de Anderson y Ball (1978); Campbell y Stanley (1966); Fitz-Gibbon y Morris (1978 y 1987); Guba y Lincoln (1981) y Walberg y Haertel (1990).

30. Para una revisión de esta literatura, pueden consultarse obras ya consolidadas como las de Guba y Lincoln (1981); Fishman (1991a; 1991b; 1992) y Shadish, Cook y Levinton (1991). Desde una perspectiva integradora de la aproximación cuantitativa y cualitativa, pueden consultarse los trabajos de Reichardt y Rallis (1994); Tashakkori y Teddlie (1998) y Russell (1999).

31. Los análisis cuantitativos son requeridos por los diseños de evaluación experimental y cuasi-experimental. La estadística empleada –descriptiva, inferencial, correlación, regresión, etc.– puede variar en su grado de sofisticación, aunque, como recuerdan los metodólogos, debería ajustarse a las características del estudio y a la formación de sus responsables. Sobre el par-

ticular, consúltense los trabajos de Kirk (1982); Kerlinger (1986); Mohr (1988); Light, Singer y Willett (1990); Pedhazur y Schmelkin (1991); Posavac y Carey (1992); Fowler (1993) y, por supuesto, Rossi y Freeman (1989).

32. De carácter inductivo, las técnicas cualitativas son esenciales para los diseños naturalistas o pluralistas. Payne (1994) fija en cuatro las variedades de análisis cualitativos: análisis fenomenológicos; análisis de contenido; inducción analítica y análisis comparativo constante. En esta línea se situarían Guba y Lincoln (1989), Patton (1980; 2001) y, sobre todo, Stake (1975; 1978; 1980; 1995), quien deja al margen el debate sobre la causalidad y defiende vivamente la importancia de los métodos cualitativos y, entre ellos, los estudios de caso, como un valioso instrumento de mejora de la gestión y generación de aprendizaje. Otras obras destacadas en la orientación cualitativa son: Stake y Gjerde (1974); Stake y Easley (1978). Scriven (1980); Shaw (1999) y Denzin y Lincoln (2000).

33. El impacto se distingue del producto y el resultado de una política o programa (Scriven, 1991).

34. Una de las grandes enseñanzas de este tipo de diseño de investigación concierne a la construcción de hipótesis de impacto. Para hablar del impacto, según ambos autores, es necesario que concurren tres condiciones: que la causa preceda al efecto; que la causa y el efecto estén correlacionados; y que la hipótesis de impacto no sea falseada por hipótesis alternativas plausibles. Para descartar estas alternativas se emplearían grupos de control.

35. Para conocer más sobre el debate de la validez desde una perspectiva clásica, es aconsejable consultar los trabajos de Carmines y Zeller (1979); Zeller (1990). Una discusión sobre las visiones contemporáneas se puede encontrar en Messick (1985).

36. Las obras de Campbell (1969; 1982; 1994), Campbell y Stanley (1966), Campbell y Cook (1976) y Cook y Campbell (1979) se convertirían en verdaderas apoloías de la evaluación experimental y de su utilidad formativa en sus diversas aplicaciones a la Administración Pública.

37. La evaluación se articularía en tres macro-actividades: la determinación de los criterios de mérito (dimensiones a través de las cuales se pretende estudiar al evaluando); la fijación de los estándares comparativos; y la medida de la actuación (*performance*) del evaluando en relación con los criterios de actuación previamente establecidos. En suma, la totalidad del proceso

debería satisfacer la necesidad de saber si el evaluando es o no eficaz (Scriven, 1980) y, desde una perspectiva sumativa, si es procedente o no su persistencia futura.

38. Con un sentido muy próximo al de la condición formulada por M.Scriven, Wholey (1983: 119) subordinaría la decisión de iniciar una evaluación a la constatación de que la utilidad proporcionada por la información a extraer supera los costes de obtención.

39. La evaluación evitaría realizarse desde la perspectiva de la gestión del programa y desde la del cliente, y se situaría en un plano superior que evitaría la influencia de factores contextuales.

40. *Vid.* las críticas realizadas por Weiss y Rein (1970) y las contracríticas de Shadish, Cook y Levinton (1991: 132-133).

41. Las dificultades de inserción de la evaluación en la teoría de la elaboración de políticas son resueltas, simplemente, a través de la "alienación" del investigador, convertido en un engranaje de la gestión orientada a los resultados.

42. Como ilustran los cuatro tipos de estudio de previsión de evaluaciones de gran amplitud, complejidad e intensidad propuestas por J.S.Wholey: análisis de evaluabilidad (*evaluability assessment*); evaluación rápida (*rapid feedback evaluation*); control de resultados (*performance monitoring*) y evaluación intensiva (*intensive evaluation*).

43. Por ejemplo, el contenido de los seis niveles en los que se articulan las actividades de evaluación prescritos por Wholey (1983: 202): identificación de objetivos y prioridades y diagnóstico de potencialidades del programa; desarrollo de sistemas para el análisis del rendimiento del programa; fijación (negociación) de niveles de mejora del rendimiento; desarrollo de sistemas que faciliten la utilización de la información generada; valoración y documentación de los resultados observados; comunicación del rendimiento de la actuación llevada a cabo.

44. Contextualización, observación, flexibilidad-adaptabilidad y comunicación son las características del planteamiento evaluativo de R.E.Stake: estudiar en profundidad las circunstancias que componen el telón de fondo de las intervenciones; esforzarse en observar las necesidades del cliente; permitir que los puntos críticos sobre los que se debe incidir afloren y evolucionen durante el curso de la evaluación; y mantener una comunicación fluida con los actores involucrados.

45. Para Wholey (1977: 42) es el pragmatismo el que orienta la construcción del saber, como en el caso de la validez de la inferencia causal.

46. En esta línea, Weiss (1973) enfatiza la idea de que los resultados de la evaluación son introducidos en un escenario de lucha política en el que, además de muchos otros elementos, operan variados intereses e ideologías.

47. Esta idea estaría presente en varias obras de Weiss (1979; 1986: 214-135; 1987: 66-69). Otra interpretación de la evaluación y la política como una única realidad puede encontrarse, igualmente, en la obra de Monnier (1995).

48. Como recordaban Feldman y March (1981), la información fundamentada empíricamente es un símbolo de racionalidad: con independencia de su utilidad en relación con los problemas en cuestión, el simple hecho de ofrecer información es una expresión importante de legitimidad y competencia organizativa que revaloriza los procesos de toma de decisiones.

49. La idea de la *policy-shaping community* pretende llamar la atención sobre el hecho de que las decisiones no se "toman", sino que se desprenden de un proceso de ajuste mutuo entre los actores intervinientes (Cronbach, et. al. 1980).

50. Al considerar el problema de la validez, Cronbach emplea la terminología "UTOS" y "utos": define los UTOS como poblaciones de unidad, tratamiento, observación y contexto, objetos de la evaluación; y "utos" (en minúscula) como una muestra del objeto, de la cual se extraen los datos. Mientras que para otros autores, por ejemplo, Campbell, la validez interna atañe a la naturaleza de las relaciones entre el tratamiento y la observación de la muestra, para Cronbach la validez interna tiene ya una connotación de generalización, concerniendo a la inferencia de la muestra al conjunto de las poblaciones. Cronbach insiste en la relación entre la muestra y la población y da la máxima prioridad a la validez externa concerniente al conjunto de poblaciones. Asimismo, la validez nunca es una característica de los métodos utilizados, sino una propiedad de las conclusiones.

51. Se trata, en suma, de un enfoque adherido a las líneas más firmemente defensoras del estudio del proceso de elaboración de políticas, en la línea defendida por G. Majone, quien ve en la persuasión y el ajuste mutuo los motores del desarrollo de las políticas.

52. Weiss (1979) y Kelly y Maynard-Moody (1993) se decantarían por una perspectiva con-

ceptual e interpretativa del trabajo de investigación, más próxima a posiciones hermenéuticas.

53. Intentando superar el pesimismo reinante entre los evaluadores a causa de los pobres resultados cosechados y la escasa utilidad de las experiencias llevadas a cabo en el contexto de los programas sociales de los sesenta (Horst, et. al., 1974), con el objetivo de ensalzar la "ciencia de la evaluación", Weiss iniciaría una verdadera cruzada que la llevaría a lo largo de sus obras de los setenta y principios de los ochenta a examinar, en primer lugar, las causas de tan empobrecido empleo de técnicas y escaso fomento de la evaluación (la determinación de las causas del mal empleo de las técnicas de evaluación y de los métodos de investigación social en las tareas orientadas a la elaboración de las políticas públicas la llevaría a poner de relieve la incompatibilidad, en términos epistemológicos, entre investigación y programación social) y, en segundo lugar, a realizar una reflexión con fines prescriptivos sobre algunos aspectos metodológicos: aun reconociendo sus limitaciones, Weiss defendería el empleo tanto de los métodos cuantitativos como de los cualitativos.

54. Lindblom (1986 y 1990).

55. Sobre el papel desempeñado por el evaluador, consúltese el trabajo de Boyle y Lemaire (1999).

REFERENCIAS BIBLIOGRÁFICAS

- ANDERSON, S.B. y BALL, S. (1978), *The Profession and Practice of Program Evaluation*. San Francisco: Jossey-Bass.
- BALLART, X. (1998), "La <<industria>> de la evaluación y bibliografía básica", *Gestión y Análisis de Políticas Públicas*, 11-12.
- BANCO MUNDIAL (1996), "Designing Project Monitoring and Evaluation", *Lessons and Practices*, 8.
- BOYLE, R. y LEMAIRE, D. (1999), *Building Effective Evaluation Capacity: Lessons from Practice*. New Brunswick: Transaction.
- BUSTELO, M. (2003), "¿Qué tiene de específico la metodología de evaluación?", en R. Bañón, *La Evaluación de la Acción y de las Políticas Públicas*. Madrid: Díaz de Santos.

- CAMPBELL, D.T. (1969), "Reforms as Experiments". *American Psychologist*, 24.
- CAMPBELL, D.T. (1982), "Experiments as arguments". *Knowledge: Creation, Diffusion, Utilization*, 3.
- CAMPBELL, D.T. (1994), "Foreword". En R.K.Yin, *Case Study Research*. Londres: Sage.
- CAMPBELL, D.T. y COOK, T.D. (1976), "The design and conduct of quasi experiments and true experiments in field setting". En M.D.Dunnette (Ed.), *Handbook of Industrial and Organizational Psychology*. Chicago: Rand McNally.
- CAMPBELL, D.T. y STANLEY, J.C. (1963), *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally.
- CARMINES, E.G. y ZELLER, R.A. (1979), *Reliability and Validity Assessment*. Newbury Park, Ca.: Sage.
- COOK, T.D. y CAMPBELL, D.T. (1979), *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Boston: Houghton Mifflin.
- CRONBACH, L.J. et. al. (1980), *Toward a Reform of Program Evaluation*. San Francisco: Jossey-Bass.
- CHELIMSKY, E. (1997), "The next transformations in evaluation". En E.Chelimsky y W.R.Shadish (Eds.), *Evaluation for the 21st Century. A Handbook*. Londres: Sage.
- CHELIMSKY, E. (Ed.) (1985), *Program Evaluation: Patterns and Directions*. Washington, D.C.: American Society for Public Administration.
- CHELIMSKY, E. y SHADISH, W.R. (Eds.) (1997), *Evaluation for the 21st Century. A Handbook*. Londres: Sage.
- CHEN, H-T. (1990), *Theory-Driven Evaluations*. Beverly Hills, California: Sage.
- DENZIN, N.K. y LINCOLN, Y.S. (2000), *Handbook of Qualitative Research* (2^a Ed.). Thousand Oaks, Ca.: Sage.
- FELDMAN, M.S. y MARCH, J.G. (1981), "Information as signal and symbol". *Administrative Science Quarterly*, 26, 1.
- FISHMAN, D.B. (1991a), "An introduction to the experimental versus the pragmatic paradigm in evaluation". *Evaluation and Program Planning*, 14.
- FISHMAN, D.B. (1991b), "The experimental versus the pragmatic paradigm: Summary and conclusions". *Evaluation and Program Planning*, 14.
- FISHMAN, D.B. (1992), "Postmodernism comes to program evaluation. A critical review of Guba and Lincoln's Fourth Generation Evaluation". *Evaluation and Program Planning*, 15.
- FITZ-GIBBON, C.T. y MORRIS, L.L. (1978), *How to Design a Program Evaluation*. Beverly Hills, California: Sage.
- FITZ-GIBBON, C.T. y MORRIS, L.L. (1987), *Evaluator's Handbook*. Beverly Hills, California: Sage.
- FONTAINE, C. y MONNIER, E. (1998), *Pratiques Françaises et Européennes*. Paris: Pouvoirs Locaux.
- FOWLER, F.J. (1993), *Survey Research Methods* (2^a Ed). Newbury Park, Ca.: Sage.
- GUBA, E.G. y LINCOLN, Y.S. (1981), *Effective Evaluation: Improving the Usefulness of Evaluation Results through Responsive and Naturalistic Approaches*. San Francisco: Jossey-Bass.
- GUBA, E.G. y LINCOLN, Y.S. (1989), *Fourth Generation Evaluation*. Londres: Sage.
- HORST, P. et.al. (1974), "Program management and the federal evaluator". *Public Administration Review*, 34.
- HUDSON, J.; MAYNE, J. y THOMLISON, R. (Eds.) (1992), *Action-Oriented Evaluation in Organizations: Canadian Practices*. Toronto: Wall & Emerson.
- KELLY, M. y MAYNARD-MOODY, S. (1993), "Policy analysis in the post-positivist era: Engaging stakeholders in evaluating the Economic Development Districts program". *Public Administration Review*, 53, 2.
- KERLINGER, F.N. (1986), *Foundations of Behavioral Research* (3^a Ed.). Nueva York: Holt, Rinehart, and Winston.
- KIRK, R.E. (1982), *Experimental Design* (2^a Ed.). Belmont, Ca: Brooks / Cole.
- LIGHT, R.J.; SINGER, J.D. y WILLETT, J.B. (1990), *By Design: Planning Research*

- on *Higher Education*. Cambridge, Harvard University.
- LINDBLOM, Ch. (1986), "Who needs what social research for policymaking?". *Knowledge: Creation, Diffusion, Utilization*, 7.
- LINDBLOM, Ch. (1990), *Inquiry and Change. The Troubled Attempt to Understand and Shape Society*. New Haven, Conn.: Yale University Press.
- LOWI, T. (1964), "American business, public policy, case studies and political theory". *World Politics*, 16 (4).
- MESSICK, S. (1985), "The once and future issues of validity: Assessing the meaning and consequences of measurement". En H.Wainer y H.I.Braun (Eds.), *Test Validity*. Hillsdale, NJ: Lawrence Earlbaum.
- MOHR, L.B. (1988), *Impact Analysis for Program Evaluation*. Chicago, Illinois: The Dorsey Press.
- MONNIER, E. (1990-91), "Objetivos y destinatarios de las evaluaciones". *Documentación Administrativa*, 224-225.
- MONNIER, E. (1995), *Evaluación de la Acción de los Poderes Públicos*. Madrid: Instituto de Estudios Fiscales.
- NACHMIAS, D. (1979), *Public Policy Evaluation: Approaches and Methods*. Nueva York: St. Martin.
- NIOCHE, J-P. (1982), "De l'évaluation a l'analyse des politiques publiques". *Revue Française de Science Politique*, 32.
- PALUMBO, D.J. (Ed.) (1987), *The Politics of Program Evaluation*. Newbury Park, California: Sage.
- PATTON, M.Q. (1980), *Qualitative Evaluation and Research Methods*. Londres: Sage.
- PATTON, M.Q. (2001), *Qualitative Research and Evaluation Methods*. Londres: Sage.
- PAYNE, D.A. (1994), *Designing Educational Project and Program Evaluations: A practical Overview Based on Research and Experience*. Boston: Kluwer.
- PEDHAZUR, E.J. y SCHMELKIN, L.P. (1991), *Measurement, Design and Analysis: An Integrated Approach*. Hillsdale, NJ: Lawrence Erlbaum, 1991.
- PETERS, G.B. (1998), "El surgimiento, declive y resurgimiento de la evaluación de políticas públicas en el Gobierno americano". *Gestión y Análisis de Políticas Públicas*, 11-12 (1998).
- POSAVAC, E.J. y CAREY, R.G. (1992), *Program Evaluation, Methods and Case Studies*. Englewood Cliffs, N.J.: Prentice Hall.
- PRESKILL, H. y TORRES, R.T. (1988), *Evaluative Inquiry for Learning in Organizations*. Thousand Oaks, Sage.
- RADAELLI, C.M. y DENTE, B. (1995), "Valutazione e scienza politica: un dialogo possibile". En G.Regonini (A cura di), *Politiche Pubbliche e Democrazia*. Naples: Edizione Scientifiche Italiane.
- REGONINI, G. (2001), *Capire le Politiche Pubbliche*. Bologna: Il Mulino.
- REICHARDT, C. y RALLIS, S. (1994), "The qualitative-quantitative debate: New perspectives". *New Directions for Program Evaluation*, 61.
- ROSSI, P.H. y FREEMAN, H.E. (1989), *Evaluation: A Systematic Approach*. Newbury Park, Ca.: Sage. [Ed. en español: *Evaluación: Un Enfoque Sistemático para Programas Sociales*. México: Trillas, 1990].
- RUSSELL, B.H. (1999), *Social Research Methods: Qualitative and Quantitative Approaches*. Thousand Oaks: Sage.
- RUTMAN, L. (Ed.) (1980), *Planning Useful Evaluations: Evaluability Assessment*. Newbury Park, California: Sage.
- SABATIER, P.A. y JENKINS-SMITH, H.C. (1993), «The dynamics of policy-oriented learning», en P.A.Sabatier y H.C.Jenkins-Smith (Eds.), *Policy Change and Learning: An Advocacy Coalition Approach*. Boulder, Colorado: Westview.
- SCRIVEN, M. (1980), *The Logic of Evaluation*. Inverness: Edgepress.
- SCRIVEN, M. (1991), *Evaluation Thesaurus* (4ª Ed.). Newbury Park, Ca.: Sage.

- SHADISH, T. COOK y LEVINTON, L. (1991), *Foundations of Program Evaluation. Theories of Practice*. Newbury Park, Ca. Sage.
- SHAW, I. (1999), *Qualitative Evaluation*. Thousand Oaks, Sage.
- STAKE, R.E. (1978), "The case study method in social inquiry". *Educational Researcher*, 7.
- STAKE, R.E. (1980), "Program evaluation. Particularly responsive evaluation". En W.B.Dckrell y D.Hamilton (Eds.), *Rethinking Educational Research*. Londres: Hodder & Stoughton.
- STAKE, R.E. (1995), *The Art of Case Study Research*. Thousand Oaks, Sage.
- STAKE, R.E. (Ed.) (1975), *Evaluating the Arts in Education: A Responsive Approach*. Columbus, Ohio: Merrill.
- STAKE, R.E. y EASLEY, J.A. (1978), *Case Studies in Science Education*. Champaign: University of Illinois.
- STAKE, R.E. y GJERDE, C. (1974), *An evaluation of TCITY (The Twin City Institute for Talented Youth)*. Chicago: Rand McNally.
- STUFFLEBEAM, D. y SHINKFIELD, A. (1995), *Evaluación Sistemática. Guía Teórica y Práctica*. Madrid: Paidós / MEC.
- SUBIRATS, J. (1994), *Policy Instruments, Public Deliberation and Evaluation Processes*. Madrid: Instituto Juan March.
- SUCHMAN, E. (1967), *Evaluative Research*. Nueva York: Russel Sage Foundation.
- TASHAKKORI, A. y TEDDLIE, C. (1998), *Mixed Methodology: Combining Qualitative And Quantitative Approaches*. Thousand Oaks, Sage.
- TOULEMONDE, J. (1995) "The emergence of an evaluation profession in European countries: the case of structural policies". *Knowledge and Policy*.
- TOULEMONDE, J.; LEEUW, F.L. y BROUWERS, A. (1999), "Evaluation activities in Europe: A quick scan of the market in 1998". *Evaluation*, 5 (4).
- TYLER, R.W.; GAGNE, R.M. y SCRIVEN, M. (Eds.), *Perspectives of Curriculum Evaluation*. Chicago: Rand McNally.
- VEDUNG, E. (1997), *Public Policy and Program Evaluation*. New Brunswick, N.J.: Transaction.
- WALBERG, H.J. y G.D.Haertel (Eds.) (1990), *The International Encyclopedia of Educational Evaluation*. Oxford: Pergamon.
- WEISS, C.H. (1973), "Where politics and evaluation meet". *Evaluation*, 1.
- WEISS, C.H. (1979), "The many meanings of research utilization". *Public Administration Review*, 39, 5.
- WEISS, C.H. (1986), "Research and policy-making: a limited partnership". En F.Heller (Ed.), *Use and Abuse of Social Science*. Londres: Sage.
- WEISS, C.H. (1987), "Postscript to "Where politics and evaluation meet". En D.Palumbo (Ed.), *The Politics of Program Evaluation*. Londres: Sage
- WEISS, C.H. (1998), *Evaluation. Methods for Studying Programs and Policies*. Upper Saddle River, N.J.: Prentice Hall.
- WEISS, R.S. y Rein, M. (1970), "The evaluation of broad-aim programs: experimental design, its difficulties and an alternative". *Administrative Science Quarterly*, 15.
- WHOLEY, J.S. (1977), "Evaluatibility Assessment". En L.Rutman (Ed.), *Evaluation Research Methods: A Basic Guide*. Londres: Sage.
- WHOLEY, J.S. (1983), *Evaluation and Effective Public Management*. Boston: Little Brown.
- WHOLEY, J.S. et.al. (1970), *Federal Evaluation Policy: Analyzing the Effects of Public Programs*. Washington, D.C.: The Urban Institute.
- WILDAVSKY, A. (1979), *Speaking Truth to Power. The Art and Craft of Policy Analysis*. Boston: Little Brown.
- ZELLER, R.A., (1990), "Validity". En Walberg, H.J. y G.D.Haertel (Eds.), *The International Encyclopedia of Educational Evaluation*. Oxford: Pergamon.